

PATENT ABSTRACTS OF JAPAN

(11)Publication number : 07-160292

(43)Date of publication of application : 23.06.1995

(51)Int.Cl. G10L 7/04

G10L 9/18

(21)Application number : 05-306898 (71)Applicant : SONY CORP

(22)Date of filing : 07.12.1993 (72)Inventor : AKAGIRI KENZO

(54) MULTILAYERED CODING DEVICE

(57)Abstract:

PURPOSE: To improve compressibility and to perform bit distribution among channels by detecting energy for each channel and performing encoding in which information quantity of each channel is allocated with a ratio in accordance with this detected output.

CONSTITUTION: This device has compression-encoding circuits 2021-202n and entropy encoding circuit 2031-203n being a first encoding means in which digital audio signals of each channel CH1-CHn are inputted and the digital audio signals are compression-encoded. This device has also a second encoding means in which an output encoded signal of the first encoding means is supplied, energy for each channel is detected, and encoding is performed in which information quantity of each channel is allocated. This second encoding means has a Log

spectrum envelop detecting circuit 208, a distribution deciding circuit 209, an adaptive quantization circuits 2031-203n, and a sub-information compression circuit 2041-204n. Thereby, bit distribution among channels and high compressibility can be realized.

LEGAL STATUS [Date of request for examination] 07.12.2000

[Date of sending the examiner's decision of rejection]

[Kind of final disposal of application other than the examiner's decision of rejection or application converted registration]

[Date of final disposal for application]

[Patent number] 3404837

[Date of registration] 07.03.2003

[Number of appeal against examiner's decision of rejection]

[Date of requesting appeal against examiner's decision of rejection]

[Date of extinction of right] 07.03.2006

* NOTICES *

JPO and INPIT are not responsible for any damages caused by the use of this translation.

1.This document has been translated by computer. So the translation may not reflect the original precisely.

2.**** shows the word which can not be translated.

3.In the drawings, any words are not translated.

CLAIMS

[Claim(s)]

[Claim 1] The multilayer coding equipment carry out [coming to have in the 2nd coding means which performs coding which the digital audio signal of each channel is inputted, and the output coded signal of the 1st coding means which carries out compression coding of the digital audio signal concerned, and the coding means of the above 1st is supplied, detects the energy for every channel in the multilayer coding equipment which carries out multilayer coding of the digital audio signal of a multiple channel, and assigns the amount of information of each channel to this detection output by the ***** ratio and] as the description.

DETAILED DESCRIPTION

[Detailed Description of the Invention]

[0001]

[Industrial Application] This invention relates to stereos, such as for example, a motion-picture film projection system, a video tape recorder, and a videodisk player, and the multilayer coding equipment which carries out compression coding of the digital audio signal of the multi-channel used in the so-called multi-surround sound system.

[0002]

[Description of the Prior Art] Although it is in the technique and the equipments of high efficiency coding of a signal, such as an audio or voice, variously For example, block the audio signal of a time domain etc. for every unit time amount, change the signal of the time-axis for this the block of every into the signal on a frequency shaft (orthogonal transformation), and it divides into two or more frequency bands. Without blocking the blocking frequency band division method encoded for every band, the so-called conversion coding (transformation coating), the audio signal of a time domain, etc. for every unit time amount Band division coding (subband coding: SBC) which is the deblocking frequency band division method divided and encoded to two

or more frequency bands can be mentioned. Moreover, the technique and equipment of high efficiency coding which combined above-mentioned band division coding and above-mentioned conversion coding are also considered, in this case, after performing band division by the above-mentioned band division coding, orthogonal transformation of the signal for this every band is carried out to the signal of a frequency domain, and coding is performed for each [by which orthogonal transformation was carried out] of this band of every.

[0003] Here, as a filter for band division of band division coding mentioned above, there are filters, such as QMF, for example and this is 1976 R.E.Crochiere. Digital coding of speech in subbands Bell Syst.Tech.J.Vol.55 It is stated to No.8 1976. moreover, ICASSP 83 BOSTON Polyphase Quadrature filters-A new subband coding technique Joseph H.Rothweiler **** -- the filter division technique and equipment of a ** bandwidth are described.

[0004] Moreover, as orthogonal transformation mentioned above, an input audio signal is blocked by predetermined unit time amount (frame), and orthogonal transformation which changes a time-axis into a frequency shaft by performing a fast Fourier transform (FFT), cosine conversion (DCT), MODIFIED DCT conversion (MDCT), etc. for this every block occurs, for example. About this MDCT, it is ICASSP 1987 Subband/Transform Coding Using Filter Bank Designs Based on Time Domain Aliasing Cancellation J.P.Princen. It is stated to A.B.Bradley Univ.of Surrey Royal Melbourne Inst.of Tech.

[0005] Furthermore, as frequency-division width of face in the case of quantizing each frequency component by which frequency band division was carried out, there is band division which took into consideration human being's acoustic-sense property, for example. That is, in the higher region currently generally called the critical band (critical band), bandwidth may divide an audio signal into the band of plurality (for example, 25 bunt) with bandwidth which becomes large. Moreover, in case the data for every band at this time are encoded, predetermined bit allocation or coding according to accommodative bit allocation the whole band is performed for every band. For example, in case the multiplier data which MDCT processing was carried out [above-mentioned] and obtained are encoded by the above-mentioned bit allocation, coding will be performed with the accommodative allocation number of bits to the MDCT multiplier data for every band obtained by MDCT processing for every above-mentioned block.

[0006] Two technique and equipment of a degree are known as the above-mentioned bit allocation technique and equipment for it. IEEE Transactions of Acoustics, Speech, and Signal Processing, vol.ASSP-25, No.4, and August 1977 Bit allocation is performed based on the magnitude of the signal for every band. Moreover, ICASSP 1980 The critical band coder--digital

encoding of the perceptual requirements of the auditory system M.A.Krassner MIT The technique and equipment which obtain the required S/N for every band and perform fixed bit allocation by using auditory masking then are described.

[0007] Here, in the high efficiency compression coding method of an audio signal using subband coding which was mentioned above, for example, the property on human being's acoustic sense is used, and a method which compresses about 1/5 of audio data is already put in practical use. In addition, as a low bit rate coding method which compresses about 1/5 of this audio data, it is the so-called ATRAC (Adaptive Transform Acoustic Coding), for example. The method called exists.

[0008] Furthermore, in stereos, such as not only the case of the usual audio equipment but for example, a motion-picture film projection system, a high definition television, a video tape recorder, a videodisk player, etc., or a multi-surround sound system, it is coming to treat the audio or sound signals of a multiple channel, such as 4-8 etc. channels, for example, and to perform high efficiency coding which reduces bit rates also in this case is desired.

[0009] The device which multi-channelization of a digital audio is progressing, for example, treats the digital audio signal of eight channels in business use especially is permeating. As a device treating the digital audio signal of the eight above-mentioned channels, there is a motion-picture film projection system etc., for example. Moreover, also in stereos, such as a high definition television, a video tape recorder, and a videodisk player, or a multi-surround sound system, it is coming to treat the audio or sound signals of a multiple channel, such as 4-8 etc. channels, for example.

[0010] Here, in the motion-picture film projection system treating the digital audio signal of the eight above-mentioned channels, recording the digital audio signal, a left channel, a left center channel, a center channel, a light center channel, a light channel, a surround left channel, a surround light channel, and a subwoofer channel, of eight channels is being performed as opposed to the above-mentioned motion-picture film. In addition, each eight above-mentioned channels recorded on the above-mentioned motion-picture film correspond with the surround light loudspeaker arranged on the left loudspeaker arranged at the screen side on which the image reproduced, for example from the image recording field of the motion-picture film concerned is projected by the projector, a left pin center, large loudspeaker, a center loudspeaker, a light center loudspeaker, a light loudspeaker, a subwoofer loudspeaker, SARAUNDORFUTOSUPIKA arranged on left-hand side so that a seat may be surrounded, and right-hand side.

[0011] However, since it is difficult for a motion-picture film to secure the field in which the 16-bit audio data by which linear quantization was carried out are [the amount of / above-mentioned / no less than eight channels] recordable with a sampling frequency of 44.1kHz which is used with the so-called CD (compact disk) etc. when recording the digital audio signal of the eight above-mentioned channels on a motion-picture film, it is made as [record / the audio data of the eight above-mentioned channels / compress and]. for example, the thing for which optimal bit assignment is performed in consideration of the property of human being's acoustic sense which was mentioned above as the compression approach which compresses the digital audio data concerned of eight channels -- for example, a sampling frequency of 44.1kHz which is recorded on the so-called CD (compact disk) etc. -- 16-bit digital audio data -- about -- though compressed into one fifth, he is trying to apply said low bit rate coding methods (the so-called ATRAC method etc.) which attain about the same tone quality as CD

[0012] Moreover, since a blemish etc. tends to generate a medium called a film on a front face, in having recorded the digital data, while it had been original, a data chip is not used violently. For this reason, the capacity of an error correction sign needs to become very important, and it is necessary to perform the above-mentioned data compression to extent recordable on the record section on the above-mentioned film also including that correction sign.

[0013]

[Problem(s) to be Solved by the Invention] However, by the method which carries out high efficiency coding of the audio signal of the above multi-channel, since compression is performed for every channel, it is hard to say that the amount of bit allocation as the whole (the amount of cutting tool allocation) is not necessarily used effectively. For example, although little bit allocation is sufficient, by other channels, more bit allocation may be needed for a certain channel. Thus, in the conventional low bit rate coding method, it is thought with looking at the amount of bit allocation (the amount of cutting tool allocation) over each channel that there is much futility. When the amount of bit allocation (or the amount of cutting tool allocation) is being especially fixed for every channel, the above futility is considered to become still more remarkable. To raise compressibility further in recent years is also desired.

[0014] Then, this invention aims at offering the multilayer coding equipment which also enables bit allocation between channels while it can raise compressibility more in view of what mentioned above.

[0015]

[Means for Solving the Problem] This invention is proposed in order to attain the

above-mentioned purpose. The multilayer coding equipment of this invention The 1st coding means which carries out multilayer coding of the digital audio signal of a multiple channel, and the digital audio signal of each channel is inputted and carries out compression coding of the digital audio signal concerned, It is characterized by coming to have the 2nd coding means which performs coding which the output coded signal of the coding means of the above 1st is supplied, and detects the energy for every channel, is a ***** ratio and assigns the amount of information of each channel to this detection output.

[0016]

[Function] According to this invention, after carrying out compression coding of the digital audio signal of each channel with the 1st coding means, bit allocation between channels and high-pressure shrinking percentage are realized by performing coding which assigns the amount of information of each channel further to the energy for every channel by the ***** ratio with the 2nd coding means to the output coded signal of this 1st coding means.

[0017]

[Example] Hereafter, it explains, referring to a drawing about the example of this invention.

[0018] The configuration of the important section of the multilayer coding equipment of this invention example is shown in drawing 1.

[0019] The multilayer coding equipment of this invention example is what carries out multilayer coding of the digital audio signal of multiple channels CH1-CHn, as shown in this drawing 1. The digital audio signal of each channels CH1-CHn is inputted. compression coding networks 2021-202n as 1st coding means which carries out compression coding of the digital audio signal concerned 2031-203n of and entropy coding networks 2031-203n of entropy coding networks of the coding means of the above 1st The coded signal (digital audio signal by which compression coding was carried out from 2021-202n of compression coding networks) of the Maine information is supplied. from -- the energy for every channels CH1-CHn The Log spectra RUEMBE rope detector 208 to detect and the allocation (bit allocation is performed) decision circuit 209 which determines the amount of information assigned to each channel by the ***** ratio as a detection output from this Log spectra RUEMBE rope detector 208, It is based on bit allocation information from this allocation decision circuit 209. 2051-205n (illustration of 2052-205n of circuits is omitted) of adaptive-quantization circuits -- setting -- the 2031-203n of the above-mentioned entropy coding networks from, while quantizing the coded signal of the Maine information accommodative 2041-204n of sub information-compression circuits It sets. The above-mentioned compression coding networks 2021-202n since -- the 2nd

coding means () which carries out compression coding of the sub information (word length information and information on a scale factor) relevant to compression coding of a digital audio signal accommodative [the Log spectra RUEMBE rope detector 208, the allocation decision circuit 209,] [adaptive-quantization circuit 2031] It is characterized by coming to have 203n and 2041-204n of sub information-compression circuits.

[0020] In addition, the digital audio signal of each channel by which compression coding was carried out with the multilayer coding equipment of this invention is transmitted through a transmission line, or is recorded on a record medium. As the record medium concerned, it is recorded on record to tape-like record media, such as disk-like record media, such as record to a motion-picture film, and an optical disk, a magneto-optic disk, a phase-change optical disk, a magnetic disk, and a magnetic tape, semiconductor memory, an IC card, etc.

[0021] Here, each above-mentioned channel ** in the case of performing record to this motion-picture film will correspond to the digital surround-sound system with which a loudspeaker is arranged as 1-CH8 is shown in drawing 2 , using a motion-picture film as the above-mentioned record medium. The channel corresponding to each loudspeaker is eight, a center (C) channel, a subwoofer (SW) channel, a left (L) channel, a left center (CL) channel, a light (R) channel, a light center (CR) channel, a left surround (LB) channel, and a light surround (RB) channel.

[0022] That is, in this drawing 2 , each channel corresponding to the above-mentioned loudspeaker arrangement corresponds with the left loudspeaker 106 arranged at the screen 101 side on which the image reproduced from the image recording field of the motion-picture film concerned is projected by the projector (projector 100), the left pin center, large loudspeaker 104, the pin center, large loudspeaker 102, the light pin center, large loudspeaker 105, the light loudspeaker 107, the surround left loudspeakers 108 and 200, the surround light loudspeakers 109 and 201, and the subwoofer loudspeaker 103.

[0023] The above-mentioned center loudspeaker 102 is arranged in the center by the side of a screen 101, outputs the playback sound by the audio data of a center channel, and outputs the most important playback sounds, such as an actor's words. The above-mentioned subwoofer loudspeaker 103 outputs the playback sound by the audio data of a subwoofer channel, outputs effectively the sound sensed as vibration rather than it calls it a sound low-pass [, such as an explosion,], and is effectively used on an explosion scene etc. in many cases. The above-mentioned left loudspeaker 106 and the light loudspeaker 107 are arranged at right and left of the above-mentioned screen 101, output the playback sound by the audio data of a left

channel, and the playback sound by the audio data of a light channel, and demonstrate the stereo sound effectiveness. The above-mentioned left center loudspeaker 104 and the light center loudspeaker 105 are arranged between the above-mentioned center loudspeaker 102, the above-mentioned left loudspeaker 106, and the light loudspeaker 107, output the playback sound by the audio data of a left center channel, and the playback sound by the audio data of a light center channel, and play the auxiliary role of the above-mentioned left loudspeaker 106 and the light loudspeaker 107, respectively. In a movie theater with it, although the normal position of an image tends to become unstable with the location of a seat, although the more real normal position of an image is made, effectiveness is demonstrated by adding the above-mentioned left center loudspeaker 104 and the light center loudspeaker 107. [in the hold number] [many / especially the screen 101 is large and] Furthermore, the above-mentioned surround left loudspeaker 108 and the surround light loudspeaker 109 are arranged so that a seat may be surrounded, they output the playback sound by the audio data of a surround left channel, and the playback sound by the audio data of a surround light channel, and are effective in giving the impression wrapped in the reverberation sound, or applause and a cheer. A thereby more three-dimensional image can be made.

[0024] It returns to drawing 1 and the multilayer coding equipment of this example is explained. this drawing 1 -- setting -- 2011-201n of each input terminal **** -- the digital audio signal of each channel which corresponds, respectively is supplied.

[0025] 2011-201n of each input terminal The digital audio signal of each supplied channel is 2021-202n of each corresponding compression coding network. It is sent and compression coding is carried out for every channel here, respectively. 2021-202n of in addition, each compression coding network About a detailed configuration, it mentions later.

[0026] 2021-202n of each compression coding network The digital audio signal by which compression coding was carried out is 2031-203n of entropy coding networks as Maine information. It is sent and variable length coding of each is carried out here. The 2021-202n of moreover, the above-mentioned compression coding networks The word length information relevant to compression coding of a digital audio signal mentioned later and the information on a scale factor are the sub information-compression circuits 2041-204n as sub information. It is sent and is compressed here based on the bit allocation information for every channel.

[0027] the 2031-203n of the above-mentioned entropy coding networks from -- each Maine information by which entropy code modulation was carried out -- 2051-205n of adaptive-quantization circuits While being sent, it is sent also to the Log spectra RUEMBE rope

detector 208. in addition -- drawing 1 -- 2032-203n of each entropy coding network of Channels CH2-CHn from -- 2102-210n of terminals with which the Maine information by which entropy code modulation was carried out corresponds, respectively from -- 2052-205n of adaptive-quantization circuits corresponding to each channels CH2-CHn which are omitting illustration while being sent -- 2152-215n of terminals It minds and is sent to the above-mentioned Log spectra RUEMBE rope detector 208.

[0028] In the Log spectra RUEMBE rope detector 208 to which each Maine information by which entropy code modulation was carried out [above-mentioned] is supplied, energy is detected from the Maine information on each [these] channel, respectively, and the Log spectra RUEMBE rope of each channel is detected. The Log spectra RUEMBE rope information on each channel concerned is sent to the above-mentioned allocation decision circuit 209.

[0029] In the allocation decision circuit 209 concerned, the channel to channel bit allocation information over the Maine information on each channel and the channel to channel bit allocation information over each sub information are determined based on the Log spectra RUEMBE rope information on each channel from the above-mentioned Log spectra RUEMBE rope detector 208.

[0030] The channel to channel bit allocation information over the above-mentioned Maine information from the above-mentioned allocation decision circuit 209 is 2051-205n of adaptive-quantization circuits. The channel to channel bit allocation information as opposed to [it is sent and] the above-mentioned sub information is the 2041-204n of the above-mentioned sub information-compression circuits. It is sent. In addition, at the example of drawing 1 , the channel to channel bit allocation information over the Maine information on Channels CH2-CHn is 2142-214n of terminals. 2052-205n of adaptive-quantization circuits which mind and are omitting illustration The channel to channel bit allocation information as opposed to [it is sent and] the sub information on Channels CH2-CHn is 2172-217n of terminals. 2162-216n of and terminals It minds and is 2042-204n of sub information-compression circuits. It is sent.

[0031] the 2051-205n of the above-mentioned adaptive-quantization circuits **** -- the channel to channel bit allocation information over the above-mentioned Maine information -- using -- the 2031-203n of each above-mentioned entropy coding network from -- the digital audio signal by which entropy code modulation was carried out is quantized accommodative.

[0032] the 2051-205n of moreover, the above-mentioned sub information-compression circuits **** -- the channel to channel bit allocation information over the above-mentioned sub information -- using -- the 2021-202n of the above-mentioned compression coding networks from

-- compression coding of the sub information (word length information, scale-factor information) is carried out.

[0033] the 2051-205n of the above-mentioned adaptive-quantization circuits 2051-205n of sub information-compression circuits from -- each output is sent to a multiplexer 206. 2052-205n of in addition, adaptive-quantization circuits which are omitting illustration of Channels CH2-CHn from -- each output -- 2122-212n of terminals it minds and sends to the above-mentioned multiplexer 206 -- having -- ** and 2052-205n of bit information-compression circuits of Channels CH2-CHn from -- each output -- 2132-213n of terminals It minds and is sent to the above-mentioned multiplexer 206.

[0034] In the multiplexer 206 concerned, the multiplexer of each supplied data is carried out, and it outputs from a terminal 207. These output data are recorded on a record medium etc. as coded data.

[0035] Moreover, as this example equipment of drawing 1 is shown, for example in drawing 3 , it is the 2021-202n of the above-mentioned compression coding networks. Receive and from the digital audio signal of each channel to supply After taking out only a high-frequency component with the high-pass filter 222 which corresponded every (every [The example of drawing 3] channels of all) at least two channels, and was prepared, respectively, Each [these] high-frequency component is added by the adder circuit 223, and processing of predetermined processing, for example, compression coding, entropy code modulation, and quantization can be performed to the addition data of the high-frequency component through a terminal 225 concerned. At this time, it is 2021-202n of compression coding networks which also take out a low-pass component from the input digital audio signal of each channel which takes out the above-mentioned high-frequency component with a low pass filter 221, respectively, and correspond through a terminal 224. It is made to send. Then, it is made to perform processing (channel to channel cross talk processing) adding the addition data which performed the above-mentioned predetermined processing to the Maine information inputted into the above-mentioned multiplexer 206 of one certain channel (the example of drawing 3 channel CH1) among each channel which takes out the above-mentioned quantity region and a low-pass component.

[0036] Here, performing adding the high-frequency component of each above-mentioned channel, and adding to the low-pass component of one certain channel is based on the following reasons.

[0037] That is, even if human being's lug has few feelings of the normal position to the component of a high region and has come out of any one loudspeaker in two or more

loudspeakers about the high-frequency component for this reason, human being has the property out of which loudspeaker it is coming, and to be hard to catch. For this reason, even if it sends each high-frequency component of the audio signal of a multiple channel only to the loudspeaker corresponding to one of channels [them], it is not sensed to human being that the high-frequency component of the audio signal of each channel is outputted only from the one loudspeaker concerned. Therefore, the high-frequency component of each channel can be compressed into one channel by adding the voice of the high-frequency component of each above-mentioned channel, and adding this addition data to the low-pass component of one certain channel as mentioned above.

[0038] 2021-202n of next, compression coding networks shown in above-mentioned drawing 1 A concrete configuration is explained.

[0039] 2021-202n of these compression coding networks Compression coding of the input digital signals, such as an audio PCM signal, is carried out using each technique of band division coding (SBC), adaptive transform coding (ATC), and adaptation bit allocation (APC-AB) then.

[0040] Hereafter, it explains using drawing 4 . 2021-202n of compression coding networks of this example shown in drawing 4 **** -- while a filter etc. divides an input digital signal into two or more frequency bands, orthogonal transformation was performed for every frequency band, and to every [in consideration of the acoustic-sense property of human being who mentions the spectrum data of the acquired frequency shaft later] so-called critical band width of face (critical band), bit allocation was carried out accommodative and it has encoded. At this time, the band which divided critical band width of face further is used in a high region. Of course, the frequency-division width of face of not blocking according to a filter etc. is good also as division-into-equal-parts ****. Furthermore, in this example, while changing a block size (block length) accommodative according to an input signal before orthogonal transformation, the block which subdivided further critical band width of face (critical band) is performing floating processing in the critical band unit or the high region. This critical band is the frequency band divided in consideration of human being's acoustic-sense property, and is a band which that noise in case the mask of the pure sound concerned is carried out by the narrow-band band noise of the same strength near the frequency of a certain pure sound has. The perimeter wave number band whose bandwidth this critical band is large like the high region, and is the above-mentioned 0-22kHz is divided into the critical band of 25.

[0041] That is, in drawing 4 , the 0-22kHz audio PCM signal is supplied to the input terminal 10. This input signal is divided into a 0-11kHz band and 11k-22kHz band by the band division filters

11, such as the so-called QMF, and, similarly the signal of a 0-11kHz band is divided into a 0-5.5kHz band and a 5.5k-11kHz band by the band division filters 12, such as the so-called QMF. The signal of 11k-22kHz band from the band division filter 11 is sent to the MDCT (Modified Discrete Cosine Transform) circuit 13 which is an example of an orthogonal transformation circuit, the signal of the 5.5k-11kHz band from the band division filter 12 is sent to the MDCT circuit 14, and MDCT processing of the signal of the 0-5.5kHz band from the band division filter 12 is carried out by being sent to the MDCT circuit 15, respectively. In addition, in each MDCT circuits 13, 14, and 15, MDCT processing is made based on the block size determined by the block decision circuits 19, 20, and 21 prepared for every band.

[0042] Here, the example of the block size in each MDCT circuits 13, 14, and 15 determined by the above-mentioned block decision circuits 19, 20, and 21 is shown in A and B of drawing 5. In addition, in B of drawing 5, the case (orthogonal transformation block size in the short mode) where an orthogonal transformation block size is short is shown, and the case (orthogonal transformation block size in long mode) where an orthogonal transformation block size is long is carried out to A of drawing 5 at it. In the example of this drawing 5, three filter outputs have two orthogonal transformation block sizes, respectively. That is, to the signal of the 0-5.5kHz band by the side of low-pass, and the signal of the 5.5k-11kHz band of a mid-range, in the case of the long block length (A of drawing 5), the measurement size in 1 block is made into 128 samples, and when a short block is chosen (B of drawing 5), the measurement size in 1 block is considered as the block for every 32 samples. On the other hand, to the signal of 11k-22kHz band by the side of a high region, in the case of the long block length (A of drawing 5), the measurement size in 1 block is made into 256 samples, and when a short block is chosen (B of drawing 5), the measurement size in 1 block is considered as the block for every 32 samples. thus -- being the same in the measurement size of an orthogonal transformation block of each band, when a short block is chosen -- carrying out -- like a high region -- time resolution -- raising -- in addition -- and the class of window used for blocking is reduced. In addition, the information which shows the block size determined in the above-mentioned block decision circuits 19, 20, and 21 is outputted from output terminals 23, 25, and 27 while it is sent to the below-mentioned adaptation bit allocation coding networks 16, 17, and 18.

[0043] Again, the spectrum data or MDCT multiplier data of a frequency domain which MDCT processing was carried out in drawing 5 in each MDCT circuits 13, 14, and 15, and was obtained is gathered for every band which divided the critical band further, and is sent to the adaptation bit allocation coding networks 16, 17, and 18 in the so-called critical band (critical band) or the

high region.

[0044] In the adaptation bit allocation coding networks 16, 17, and 18, it is made to carry out re-quantization (for it to normalize and quantize) of each spectrum data (or MDCT multiplier data) according to the number of bits which divided the critical band further and which was assigned for every band in the information and critical band (critical band), or high region of the above-mentioned block size. The data encoded by the adaptation bit allocation coding networks 16, 17, and 18 are taken out through output terminals 22, 24, and 26. Moreover, in the adaptation bit allocation coding networks 16, 17, and 18 concerned, the scale-factor information which shows whether the normalization about the magnitude of what kind of signal was made, and the bit length information (word length information) which shows by what kind of bit length quantization was carried out are also searched for, and these are also outputted to coincidence from output terminals 22, 24, and 26.

[0045] Moreover, from the output of each MDCT circuits 13, 14, and 15 in drawing 5, the energy for every band which divided the critical band further is searched for in the above-mentioned critical band (critical band) or a high region by calculating the square root of the root mean square of each amplitude value within the band concerned, for example etc. Of course, you may make it use for future bit allocation of the above-mentioned scale factor itself. In this case, since the operation of new energy count becomes unnecessary, it becomes saving of a hard scale. Moreover, it is also possible to use the peak value of amplitude value, the average, etc. instead of the energy for every band.

[0046] Here, drawing 6 explains the more concrete configuration of the above-mentioned adaptation bit allocation circuits 16, 17, and 18. In the adaptation bit allocation circuit shown in this drawing 6, the magnitude of a MDCT multiplier is called for for every block, and that MDCT multiplier is supplied to an input terminal 801. The MDCT multiplier supplied to the input terminal 801 concerned is given to the energy calculation circuit 803 for every band. In the energy calculation circuit 803 for every band, the signal energy about each band which re-divided the critical band further in the critical band or the high region is computed. The energy about each band computed in the energy calculation circuit 803 for every band is supplied to the energy dependence bit allocation circuit 804.

[0047] Bit allocation which uses the usable total bit from the usable total bit generating circuit 802 in the energy dependence bit allocation circuit 804, uses a certain rate of the 128Kbps (this example 100Kbps(es)) in this example, and makes white quantizing noise is performed. The rate that this amount of bits occupies to the above-mentioned 128Kbps increases, so that unevenness

of the spectrum of an input signal is so large that the toe nullity of an input signal is high at this time. In addition, in order to detect unevenness of the spectrum of an input signal, the sum of the absolute value of the difference of the block floating multiplier of an adjoining block is used as an index. And bit allocation proportional to the opposite numeric value of the energy of each band is performed about the calculated usable amount of bits.

[0048] The bit allocation calculation circuit 805 depending on acoustic-sense permission noise level calculates the amount of permission noises for every critical band in consideration of the so-called masking effect etc. based on the spectrum data first divided for every above-mentioned critical band, and a part for the bit which lengthened the energy dependence bit from the usable total bit so that an acoustic-sense permissible noise spectrum might next be given is distributed. Thus, the called-for energy dependence bit and the bit depending on acoustic-sense permission noise level are added, and he is trying to re-quantize each spectrum data (or MDCT multiplier data) according to the number of bits assigned to the band which divided the critical band into two or more bands further in every critical band and the high region by the adaptation bit allocation coding networks 16, 17, and 18 of drawing 4 . Thus, the encoded data are taken out through the output terminals 22, 24, and 26 of drawing 4 .

[0049] If the acoustic-sense permissible noise spectrum calculation circuit in the bit allocation circuit 805 of the above-mentioned acoustic-sense permissible noise spectrum dependence is explained in more detail, the MDCT multiplier obtained in the MDCT circuits 13, 14, and 15 will be given to the above-mentioned permissible noise calculation circuit.

[0050] Drawing 7 is the block circuit diagram showing the outline configuration of one example in which the above-mentioned permissible noise calculation circuit was explained collectively. In this drawing 7 , the spectrum data of the frequency domain from the MDCT circuits 13, 14, and 15 are supplied to the input terminal 521.

[0051] The input data of this frequency domain is sent to the energy calculation circuit 522 for every band, and when the energy of every above-mentioned critical band (critical band) calculates total of each amplitude value square within the band concerned, for example, it is called for. The peak value of amplitude value, the average, etc. may be used instead of the energy for every band of this. Generally as an output from this energy calculation circuit 522, the spectrum of the total value of for example, each band is called the bark spectrum. Drawing 8 shows the bark spectrum SB for such every critical band. However, in this drawing 8 , in order to simplify illustration, 12 bands (B1 ·B12) are expressing the number of bands of the above-mentioned critical band.

[0052] Reefing (convolution) processing which hangs and adds a predetermined weighting function to this bark spectrum SB here in order to take into consideration the effect in the so-called masking of the above-mentioned bark spectrum SB is performed. For this reason, each value of the output SB of the energy calculation circuit 522 for every above-mentioned band, i.e., this bark spectrum, is sent to the reefing filter circuit 523. This reefing filter circuit 523 consists of two or more delay elements which carry out sequential delay of the input data, two or more multipliers (for example, 25 multipliers corresponding to each band) which carry out the multiplication of the filter coefficient (weighting function) to an output from these delay elements, and a total adder which takes total of each multiplier output. In addition, the above-mentioned masking says the phenomenon which the mask of other signals is carried out by a certain signal, and stops being able to hear with the property on human being's acoustic sense, and there are a time-axis masking effect by the audio signal of a time domain and this time-of-day masking effect by the signal of a frequency domain as this masking effect. According to these masking effects, even if a noise is in the part masked, this noise can be heard. For this reason, let the noise within the limits of [which is masked] this be a permissible noise in an actual audio signal.

[0053] If one example of the multiplication multiplier (filter coefficient) of each multiplier of the above-mentioned reefing filter circuit 523 is shown, when setting the multiplier of the multiplier M corresponding to the band of arbitration to 1 here, A multiplier 0.15 with a multiplier M-2 with a multiplier M-1 a multiplier 0.0019 In a multiplier 0.0000086, reefing processing of the above-mentioned bark spectrum SB is performed by the multiplier M-3 with a multiplier M+1 by carrying out [a multiplier 0.4] the multiplication of the multiplier 0.007 for a multiplier 0.06 to the output of each delay element with a multiplier M+3 with a multiplier M+2. However, M is the integer of the arbitration of 1-25.

[0054] Next, the output of the above-mentioned reefing filter circuit 523 is sent to the subtraction machine 524. This subtraction machine 524 asks for the level alpha corresponding to the noise level in which the permission in the collapsed field mentioned later is possible the account of a top. In addition, the level alpha corresponding to the noise level (permission noise level) in which the permission concerned is possible is the level which turns into a permission noise level for every band of a critical band by performing reverse convolution processing, as mentioned later. Here, the admissible function (function expressing masking level) of the ** sake which asks for the above-mentioned level alpha is supplied to the above-mentioned subtraction machine 524. The above-mentioned level alpha is controlled by making this admissible function fluctuate. The

admissible function concerned is supplied from the function generating circuit (n·ai) 525 which is explained below.

[0055] That is, if the number given sequentially from low-pass [of the band of a critical band] is set to i, it can ask for the level alpha corresponding to a permission noise level by the following formula.

$$\text{Alpha} = S \cdot (n \cdot a^i)$$

In this formula, n and a are the reinforcement of the bark spectrum with which reefing processing of $a > 0$ and the S was carried out by the constant, and the inside (n·ai) of a formula serves as an admissible function. $n = 38$ and $a = 0.5$ can be used as an example.

[0056] Thus, the above-mentioned level alpha is called for and this data is transmitted to a divider 526. It is for carrying out the reverse convolution of the above-mentioned level alpha in the field by which the reefing was carried out [above-mentioned] in the divider 526 concerned. Therefore, Massu Kings RESSHORUDO comes to be obtained from the above-mentioned level alpha by performing this reverse convolution processing. That is, this Massu Kings RESSHORUDO serves as a permission noise spectrum. In addition, although the above-mentioned reverse convolution processing needs a complicated operation, it is performing the reverse convolution using the simplified divider 526 in this example.

[0057] Next, above-mentioned Massu Kings RESSHORUDO is transmitted to a subtractor 528 through the synthetic circuit 527. Here, the bark spectrum SB outputted namely, mentioned above from the energy detector 522 for every above-mentioned band is supplied to the subtractor 528 concerned through the delay circuit 529. Therefore, as shown in drawing 3 by the subtraction operation of above-mentioned Massu Kings RESSHORUDO and the bark spectrum SB being performed with this subtractor 528, below the level that shows the above-mentioned bark spectrum SB on the level of this masking SURESSHORUDO MS will be masked. In addition, the delay circuit 529 is formed in order to delay the bark spectrum SB from the energy detector 522 in consideration of the amount of delay in each circuit before the above-mentioned synthetic circuit 527.

[0058] The output from the subtractor 528 concerned is sent to ROM with which it was taken out through the output terminal 531, for example, allocation number-of-bits information was beforehand remembered to be through the permissible noise amendment circuit 530 (not shown). This ROM etc. outputs the allocation number-of-bits information for every band according to the output (level of the difference of the energy of each above-mentioned band, and the output of the above-mentioned noise level setting means) obtained from the above-mentioned subtractor

circuit 528 through the permissible noise amendment circuit 530.

[0059] Thus, an energy dependence bit and the bit depending on acoustic-sense permission noise level are added, and coding is performed by the adaptation bit allocation coding networks 16, 17, and 18 concerned using the allocation number-of-bits information.

[0060] That is, if it summarizes, at the adaptation bit allocation coding networks 16, 17, and 18, the spectrum data for every above-mentioned band will be quantized with the number of bits distributed according to the level of the difference of the energy of a band or peak value which divided the critical band into two or more bands further in each band band (critical band) of every and high region of the above-mentioned critical band, and the output of the above-mentioned noise level setting means.

[0061] By the way, in the case of composition, the data in which the so-called minimum audible curve RC which is human being's acoustic-sense property as shown in drawing 9 in the synthetic circuit 527 mentioned above supplied from the minimum audible curve generating circuit 532 is shown, and above-mentioned Massu Kings RESSHORUDO MS are compoundable. In this minimum audible curve, if noise absolute level becomes below this minimum audible curve, this noise can be heard. what is different by the difference in the playback volume at the time of playback even if this minimum audible curve has the same coding, ****, and a realistic digital system -- the sound for example, to a 16-bit dynamic range -- since there is no difference in the easy method of entering so much, supposing the quantizing noise of the frequency band near 4kHz a lug is the easiest to hear cannot be heard, for example, in other frequency bands, it will be thought that the quantizing noise below the level of this minimum audible curve cannot be heard. Therefore, it assumes that usage depending on which the noise near 4kHz of the dynamic range which a system has in this way cannot be heard is carried out, and if a permission noise level is obtained by compounding both this minimum audible curve RC and Massu Kings RESSHORUDO MS, the permission noise level in this case can be carried out to to the part shown with the slash in drawing 9 . In addition, in this example, the level of 4kHz of the above-mentioned minimum audible curve is doubled with the minimum level of 20 bits. Moreover, this drawing 9 shows the signal spectrum SS to coincidence.

[0062] Moreover, in the above-mentioned permissible noise amendment circuit 530, the permission [in / based on the information on an equal loudness curve for example / the output from the above-mentioned subtractor 528] noise level sent from the amendment information output circuit 533 is amended. It connected with the curve in quest of the sound pressure of the sound in each frequency which it is here, and an equal loudness curve is a characteristic curve

about human being's acoustic-sense property, for example, the same magnitude as the pure sound of 1kHz hears, and is also called the equal loudness contour of loudness rating. moreover, the minimum audible curve RC which showed this equal loudness contour to drawing 9 and abbreviation -- the same curve is drawn. In this equal loudness contour, for example near 4kHz, the magnitude as 1kHz also with the same bottom of 8-10dB hears sound pressure from the place of 1kHz, and conversely, near 50Hz, unless it is higher than the sound pressure in 1kHz about 15dB, the same magnitude does not hear. For this reason, the noise (permission noise level) beyond the level of the above-mentioned minimum audible curve is understood are good to have the frequency characteristics given in the curve according to this equal loudness contour. Since it is such, it turns out that amending the above-mentioned permission noise level in consideration of the above-mentioned equal loudness contour conforms to human being's acoustic-sense property.

[0063] The spectrum configuration depending on the acoustic-sense permission noise level described above is built with the bit allocation using a certain rate of the usable total bit 128Kbps. This rate decreases, so that the toe nullity of an input signal becomes high.

[0064] Next, the amount division technique of bits is explained. It returns to drawing 6, and the signal from the input terminal 801 with which a MDCT circuit output is supplied is also given to the smoothness calculation circuit 808 of a spectrum, and the smoothness of a spectrum is computed here. In this example, the value which broke the sum of the absolute value of the difference between the neighbors of the absolute value of a signal spectrum by the sum of the absolute value of a signal spectrum is computed as smoothness of the above-mentioned spectrum.

[0065] The output of the smoothness calculation circuit 808 of the above-mentioned spectrum is given to the rate decision circuit 809 of bit division, and the rate of bit division between bit allocation of energy dependence and the bit allocation by the acoustic-sense permissible noise spectrum is determined here. It thinks that the rate of bit division does not have the smoothness of a spectrum, so that the output value of the smoothness calculation circuit 808 of a spectrum is large, and bit allocation with emphasis on the bit allocation by the acoustic-sense permissible noise spectrum is performed rather than bit allocation of energy dependence. The rate decision circuit 809 of bit division sends a control output to the multipliers 811 and 812 which control the magnitude of bit allocation of energy dependence, and the bit allocation by the acoustic-sense permissible noise spectrum, respectively. Here, a spectrum is temporarily smooth, and when the output of the rate decision circuit 809 of bit division to a multiplier 811 takes the value of 0.8 so that weight may be set to bit allocation of energy dependence, the output of the rate decision

circuit 809 of bit division to a multiplier 812 is set to $1-0.8=0.2$. The output of these two multipliers is added by the adder 806, serves as final bit allocation information, and is outputted from an output terminal 807.

[0066] Next, the configuration of the elongation decryption circuit corresponding to the compression coding network of above-mentioned drawing 4 is shown in drawing 10.

[0067] In this drawing 10, the MDCT multiplier by which each band was quantized is given to input terminals 122, 124, and 126, and the used block-size information is given to input terminals 123, 125, and 127. In the decryption circuits 116, 117, and 118, bit allocation is canceled using adaptation bit allocation information.

[0068] Next, the signal of a frequency domain is changed into the signal of a time domain in the IMDCT circuits 113, 114, and 115. The time domain signal of these partial bands is decrypted by the IQMF circuits 112 and 111 at a whole region signal. Then, band composition is made in the IQMF circuit 111,112, and it is outputted from an output terminal 110.

[0069]

[Effect of the Invention] As mentioned above, after carrying out compression coding of the digital audio signal of each channel with the 1st coding means in the multilayer coding equipment of this invention, bit allocation between channels and high-pressure shrinking percentage are made possible by performing coding which is a ***** ratio and assigns the amount of information of each channel further to the energy for every channel with the 2nd coding means to the output coded signal of this 1st coding means.

DESCRIPTION OF DRAWINGS

[Brief Description of the Drawings]

[Drawing 1] It is the block circuit diagram showing the outline configuration of the multilayer

coding equipment of this invention example.

[Drawing 2] It is drawing for explaining arrangement of the loudspeaker in an eight-channel digital surround-sound system.

[Drawing 3] It is the block circuit diagram showing the outline configuration of the main components which perform a channel to channel cross talk.

[Drawing 4] It is the block circuit diagram showing the example of a configuration of the compression coding network of this example.

[Drawing 5] It is drawing showing the frequency and time sharing of a signal in a compression coding network.

[Drawing 6] It is the block circuit diagram showing the configuration which realizes the bit allocation technique using two persons of the magnitude of an information signal, and an acoustic-sense permissible noise spectrum.

[Drawing 7] It is the block circuit diagram showing the configuration which asks for permission noise level.

[Drawing 8] It is drawing showing the example of the masking threshold by the signal level of each band.

[Drawing 9] It is drawing showing an information spectrum, a masking threshold, and lower limit of hearing.

[Drawing 10] It is the block circuit diagram showing the example of a configuration of the elongation decryption circuit corresponding to the compression coding network of this example.

[Description of Notations]

2021 -202N ... Compression Coding Network

2031 -203N ... Entropy Coding Network

2041 -204N ... Sub Information-Compression Circuit

2051 -205N ... Adaptive-Quantization Circuit

206 ... Multiplexer

208 ... Log spectra RUEMBE rope detector

209 ... Distribution decision circuit

221 ... Low pass filter

222 ... High-pass filter

223 ... Adder circuit

11 12 Band division filter

13, 14, 15 MDCT circuit

16, 17, 18 Adaptation bit allocation coding network
 19, 20, 21 Block-size decision circuit
 22, 24, 26 Coding output terminal
 23, 25, 27 Block-size information output terminal
 116,117,118 ... Adaptation bit allocation decryption circuit
 113,114,115 ... IMDCT circuit
 112,111 ... IQMF circuit
 520 ... Permissible noise calculation circuit
 521 ... Permissible noise calculation circuit input terminal
 522 ... Energy detector for every band
 523 ... Reeking filter circuit
 524 ... Subtraction machine
 525 ... $n^{-\alpha}$ function generating circuit
 526 ... Divider
 527 ... Synthetic circuit
 528 ... Subtractor
 530 ... Permissible noise amendment circuit
 532 ... The minimum audible curve generating circuit
 533 ... Amendment information output circuit
 801 ... MDCT circuit output terminal
 802 ... The usable total bit generating circuit
 803 ... Energy calculation circuit for every band
 804 ... Bit allocation circuit of energy dependence
 805 ... Bit allocation circuit of acoustic-sense permission noise level dependence
 806 ... Adder
 807 ... Bit allotment output terminal of each band
 808 ... Smoothness calculation circuit of a spectrum
 809 ... Rate decision circuit of bit division
 811 812 ... Multiplier

(51) Int.Cl.⁶G 1 0 L 7/04
9/18

識別記号

G
C

庁内整理番号

F I

技術表示箇所

審査請求 未請求 請求項の数 1 O L (全 13 頁)

(21) 出願番号 特願平5-306898

(22) 出願日 平成5年(1993)12月7日

(71) 出願人 000002185

ソニー株式会社

東京都品川区北品川6丁目7番35号

(72) 発明者 赤桐 健三

東京都品川区北品川6丁目7番35号 ソニー株式会社内

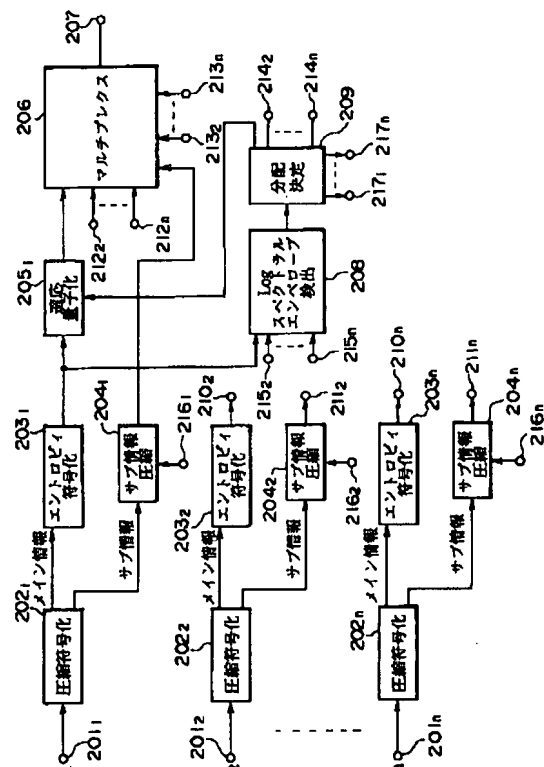
(74) 代理人 弁理士 小池 晃 (外2名)

(54) 【発明の名称】 多層符号化装置

(57) 【要約】

【構成】 各チャンネルCH1～CHnのデジタルオーディオ信号を圧縮符号化する圧縮符号化回路202₁～202_n及びエントロピ符号化回路203₁～203_nと、メイン情報(圧縮符号化されたデジタルオーディオ信号)から各チャンネル毎のエネルギーを検出してこの検出結果に応じた比率で各チャンネルのビット配分を決定するLogスペクトラルエンベロープ検出回路208及び配分決定回路209と、この配分決定回路209からのビット配分情報に基づいて、適応量子化回路205₁～205_nにおいてメイン情報を適応的に量子化すると共に、サブ情報圧縮回路204₁～204_nにおいてサブ情報(語長情報やスケールファクタの情報)を圧縮する。

【効果】 チャンネル間ビット配分と高圧縮とを実現する。



【特許請求の範囲】

【請求項1】 複数チャネルのデジタルオーディオ信号を多層符号化する多層符号化装置において、各チャネルのデジタルオーディオ信号が入力され、当該デジタルオーディオ信号を圧縮符号化する第1の符号化手段と、

上記第1の符号化手段の出力符号化信号が供給され、チャネル毎のエネルギーを検出し、この検出出力に略応じた比率で、各チャネルの情報量を割り当てる符号化を行う第2の符号化手段とを有してなることを特徴とする多層符号化装置。

【発明の詳細な説明】

【0001】

【産業上の利用分野】 本発明は、例えば、映画フィルム映写システム、ビデオテープレコーダ、ビデオディスクプレーヤ等のステレオや、いわゆるマルチサラウンド音響システムにおいて用いられるマルチチャネルのデジタルオーディオ信号を圧縮符号化する多層符号化装置に関するものである。

【0002】

【従来の技術】 オーディオ或いは音声等の信号の高効率符号化の手法及び装置には種々あるが、例えば、時間領域のオーディオ信号等を単位時間毎にブロック化してこのブロック毎の時間軸の信号を周波数軸上の信号に変換（直交変換）して複数の周波数帯域に分割し、各帯域毎に符号化するブロック化周波数帯域分割方式、いわゆる変換符号化（トランスフォームコーディング）や、時間領域のオーディオ信号等を単位時間毎にブロック化しないで、複数の周波数帯域に分割して符号化する非ブロック化周波数帯域分割方式である帯域分割符号化（サブバンドコーディング：SBC）等を挙げることができる。また、上述の帯域分割符号化と変換符号化とを組み合わせた高効率符号化の手法及び装置も考えられており、この場合には、例えば、上記帯域分割符号化で帯域分割を行った後、該各帯域毎の信号を周波数領域の信号に直交変換し、この直交変換された各帯域毎に符号化が施される。

【0003】 ここで、上述した帯域分割符号化の帯域分割用フィルタとしては、例えばQMF等のフィルタがあり、これは1976 R.E.Crochiere Digital coding of speech in subbands Bell Syst.Tech. J. Vol.55, No.8 1976に、述べられている。また、ICASSP 83, BOSTON Polyphase Quadrature filters-A new subband coding technique Joseph H. Rothweiler には等バンド幅のフィルタ分割手法及び装置が述べられている。

【0004】 また、上述した直交変換としては、例えば、入力オーディオ信号を所定単位時間（フレーム）でブロック化し、該ブロック毎に高速フーリエ変換（FFT）、コサイン変換（DCT）、モディファイドDCT変換（MDCT）などを行うことで時間軸を周波数軸に

変換するような直交変換がある。このMDCTについては、ICASSP 1987 Subband/Transform Coding Using Filter Bank Designs Based on Time Domain Aliasing Cancellation J.P.Princen A.B.Bradley Univ. of Surrey Royal Melbourne Inst.of Tech.に述べられている。

【0005】 更に、周波数帯域分割された各周波数成分を量子化する場合の周波数分割幅としては、例えば人間の聴覚特性を考慮した帯域分割がある。すなわち、一般に臨界帯域（クリティカルバンド）と呼ばれている高域程帯域幅が広くなるような帯域幅で、オーディオ信号を複数（例えば25バンド）の帯域に分割することがある。また、この時の各帯域毎のデータを符号化する際には、各帯域毎に所定のビット配分或いは、各帯域毎に適応的なビット配分による符号化が行われる。例えば、上記MDCT処理されて得られた係数データを上記ビット配分によって符号化する際には、上記各ブロック毎のMDCT処理により得られる各帯域毎のMDCT係数データに対して、適応的な配分ビット数で符号化が行われることになる。

【0006】 上記ビット配分手法及びそのための装置としては、次の2手法及び装置が知られている。IEEE Transactions of Acoustics, Speech, and Signal Processing, vol. ASSP-25, No.4, August 1977 では、各帯域毎の信号の大きさをもとに、ビット配分を行っている。また、ICASSP 1980 The critical band coder--digital encoding of the perceptual requirements of the auditory system M.A. Krasner MIT では、聴覚マスキングを利用することで、各帯域毎に必要な信号対雑音比を得て固定的なビット配分を行う手法及び装置が述べられている。

【0007】 ここで、例えば上述したようなサブバンドコーディング等を用いたオーディオ信号の高効率圧縮符号化方式においては、人間の聴覚上の特性を利用し、オーディオデータを約1/5に圧縮するような方式が既に実用化されている。なお、このオーディオデータを約1/5に圧縮する高効率符号化方式としては、例えばいわゆるATRAC（Adaptive Transform Acoustic Coding）と呼ばれる方式が存在する。

【0008】 さらに、通常のオーディオ機器の場合のみならず、例えば映画フィルム映写システム、高品位テレビジョン、ビデオテープレコーダ、ビデオディスクプレーヤ等のステレオないしはマルチサラウンド音響システムにおいては、例えば4～8チャネル等の複数チャネルのオーディオ或いは音声信号を扱うようになりつつあり、この場合においても、ビットレートを削減する高効率符号化を行うことが望まれている。

【0009】 特に、業務用においては、デジタルオーディオのマルチチャネル化が進んでおり、例えば8チャネルのデジタルオーディオ信号を扱う機器が浸透してきて、マルチチャネルのデジタルオーディオ信号の

を扱う機器としては、例えば映画フィルム映写システム等がある。また、高品位テレビジョン、ビデオテープレコーダ、ビデオディスクプレーヤ等のステレオないしはマルチサラウンド音響システムにおいても、例えば4〜8チャンネル等の複数チャンネルのオーディオ或いは音声信号を扱うようになりつつある。

【0010】ここで、上記8チャンネルのデジタルオーディオ信号を扱う映画フィルム映写システムにおいては、上記映画フィルムに対して、例えばレフトチャンネル、レフトセンタチャンネル、センタチャンネル、ライトセンタチャンネル、ライトチャンネル、サラウンドレフトチャンネル、サラウンドライトチャンネル、サブウーファチャンネルの8チャンネルのデジタルオーディオ信号を記録することが行われつつある。なお、上記映画フィルムに記録する上記8チャンネルの各チャンネルは、例えば当該映画フィルムの画像記録領域から再生された画像が映写機によって投影されるスクリーン側に配置されるレフトスピーカ、レフトセンタスピーカ、センタスピーカ、ライトセンタスピーカ、ライトスピーカ、サブウーファスピーカ、観客席を取り囲むように左側に配置されるサラウンドレフトスピーカ及び右側に配置されるサラウンドライトスピーカと対応するものである。

【0011】ただし、映画フィルムに上記8チャンネルのデジタルオーディオ信号を記録する場合において、映画フィルムには、例えばいわゆるCD（コンパクトディスク）などで用いているようなサンプリング周波数44.1kHzで16ビットの直線量子化されたオーディオデータを上記8チャンネル分も記録できる領域を確保することは困難であるため、上記8チャンネルのオーディオデータを圧縮して記録するようになされる。例えば、当該8チャンネルのデジタルオーディオデータを圧縮する圧縮方法としては、上述したような人間の聴覚の特性を考慮して最適なビット割り当てを行うことによって、例えばいわゆるCD（コンパクトディスク）などに記録されるようなサンプリング周波数44.1kHzで16ビットのデジタルオーディオデータを約1/5に圧縮しながらも、CD並の音質を達成する前記高能率符号化方式（いわゆるATRAC方式など）を適用するようにしている。

【0012】また、フィルムという媒体は、表面に傷などが発生しやすいため、デジタルデータをオリジナルのまま記録していたのでは、データ欠けが激しく実用にならない。このため、エラー訂正符号の能力が非常に重要になり、上記データ圧縮は、その訂正符号も含めて上記フィルム上の記録領域に記録可能な程度まで行う必要がある。

【0013】

【発明が解決しようとする課題】しかし、上述のようなマルチチャンネルのオーディオ信号を高能率符号化する方法では、各チャンネル毎に圧縮が行われるため、全体のレ

てのビット配分量（バイト配分量）が必ずしも有効に使用されているとは言い難い。例えば、あるチャンネルには少ないビット配分でもよいが、他のチャンネルではより多くのビット配分を必要とする場合がある。このように、従来の高能率符号化方式では、各チャンネルにわたってビット配分量（バイト配分量）を眺めると無駄が多いと思われる。特に、各チャンネル毎にビット配分量（或いはバイト配分量）が固定されている場合には、上記のような無駄がさらに顕著になると考えられる。近年は、さらに圧縮率を高めることも望まれている。

【0014】そこで、本発明は、上述したようなことに鑑み、より圧縮率を高めることができると共に、チャンネル間のビット配分をも可能とする多層符号化装置を提供することを目的としている。

【0015】

【課題を解決するための手段】本発明は、上述の目的を達成するために提案されたものであり、本発明の多層符号化装置は、複数チャンネルのデジタルオーディオ信号を多層符号化するものであり、各チャンネルのデジタルオーディオ信号が入力され、当該デジタルオーディオ信号を圧縮符号化する第1の符号化手段と、上記第1の符号化手段の出力符号化信号が供給され、チャンネル毎のエネルギーを検出し、この検出出力に略応じた比率で、各チャンネルの情報量を割り当てる符号化を行う第2の符号化手段とを有してなることを特徴とするものである。

【0016】

【作用】本発明によれば、第1の符号化手段で各チャンネルのデジタルオーディオ信号を圧縮符号化した後、この第1の符号化手段の出力符号化信号に対して、さらに第2の符号化手段によってチャンネル毎のエネルギーに略応じた比率で、各チャンネルの情報量を割り当てる符号化を行うことでチャンネル間のビット配分と高圧縮率を実現している。

【0017】

【実施例】以下、本発明の実施例について図面を参照しながら説明する。

【0018】図1には、本発明実施例の多層符号化装置の要部の構成を示す。

【0019】本発明実施例の多層符号化装置は、この図1に示すように、複数チャンネルCH1〜CHnのデジタルオーディオ信号を多層符号化するものであり、各チャンネルCH1〜CHnのデジタルオーディオ信号が入力され、当該デジタルオーディオ信号を圧縮符号化する第1の符号化手段としての圧縮符号化回路202₁〜202_n及びエントロピ符号化回路203₁〜203_nと、上記第1の符号化手段のエントロピ符号化回路203₁〜203_nからのメイン情報の符号化信号（圧縮符号化回路202₁〜202_nからの圧縮符号化されたデジタルオーディオ信号）が供給されてチャンネルCH1〜CHn毎のエネルギーを検出するエネルギー検出部204とを有する。

ルエンベロープ検出回路208と、このLogスペクトラルエンベロープ検出回路208からの検出出力に略応じた比率で各チャンネルに割り当てる情報量を決定する

(ビット配分を行う)配分決定回路209と、この配分決定回路209からのビット配分情報に基づいて、適応量子化回路205₁~205_n(回路205₂~205_nの図示は省略)において上記エントロピ符号化回路203₁~203_nからのメイン情報の符号化信号を適応的に量子化すると共に、サブ情報圧縮回路204₁~204_nにおいて上記圧縮符号化回路202₁~202_nからのデジタルオーディオ信号の圧縮符号化に関連するサブ情報(語長情報やスケールファクタの情報)を適応的に圧縮符号化する第2の符号化手段(Logスペクトラルエンベロープ検出回路208, 配分決定回路209, 適応量子化回路203₁~203_n, サブ情報圧縮回路204₁~204_n)とを有してなることを特徴とするものである。

【0020】なお、本発明の多層符号化装置によって圧縮符号化された各チャンネルのデジタルオーディオ信号は、伝送路を介して伝送されたり、記録媒体に記録されたりする。当該記録媒体として、例えば映画フィルムへの記録や、光ディスク、光磁気ディスク、相変化型光ディスク、磁気ディスク等のディスク状記録媒体、磁気テープ等のテープ状記録媒体への記録、半導体メモリ、ICカードなどへ記録される。

【0021】ここで、上記記録媒体として映画フィルムを用い、この映画フィルムへの記録を行う場合の上記各チャンネルすは1~CH8は、例えば図2に示すようにスピーカが配置されるデジタルサラウンドシステムに対応することになる。各スピーカに対応するチャンネルは、センタ(C)チャンネル、サブウーファ(SW)チャンネル、レフト(L)チャンネル、レフトセンタ(CL)チャンネル、ライト(R)チャンネル、ライトセンタ(CR)チャンネル、レフトサラウンド(LB)チャンネル、ライトサラウンド(RB)チャンネルの8つである。

【0022】すなわちこの図2において、上記スピーカ配置に対応する各チャンネルは、例えば当該映画フィルムの画像記録領域から再生された画像が映写機(プロジェクタ100)によって投影されるスクリーン101側に配置されたレフトスピーカ106、レフトセンタスピーカ104、センタスピーカ102、ライトセンタスピーカ105、ライトスピーカ107、サラウンドレフトスピーカ108及び200、サラウンドライトスピーカ109及び201、サブウーファスピーカ103と対応するものである。

【0023】上記センタスピーカ102は、スクリーン101側の中央に配置され、センタチャンネルのオーディオデータによる再生音を出力するもので例えば俳優のせりふ等の最も重要な再生音を出力する。上記サブウー

フデータによる再生音を出力するもので、例えば爆発音などの低域の音というよりは振動として感じられる音を効果的に出力するものであり、爆発シーンなどに効果的に使用されることが多いものである。上記レフトスピーカ106及びライトスピーカ107は、上記スクリーン101の左右に配置され、レフトチャンネルのオーディオデータによる再生音とライトチャンネルのオーディオデータによる再生音を出力するもので、ステレオ音響効果を発揮する。上記レフトセンタスピーカ104とライトセンタスピーカ105は、上記センタスピーカ102と上記レフトスピーカ106及びライトスピーカ107との間に配置され、レフトセンタチャンネルのオーディオデータによる再生音とライトセンタチャンネルのオーディオデータによる再生音を出力するもので、それぞれ上記レフトスピーカ106及びライトスピーカ107の補助的な役割を果たす。特にスクリーン101が大きく収容人数の多い映画館等では、座席の位置によって音像の定位が不安定になりやすいが、上記レフトセンタスピーカ104とライトセンタスピーカ107を付加することにより、音像のよりリアルな定位を作り出すのに効果を発揮する。さらに、上記サラウンドレフトスピーカ108とサラウンドライトスピーカ109は、観客席を取り囲むように配置され、サラウンドレフトチャンネルのオーディオデータによる再生音とサラウンドライトチャンネルのオーディオデータによる再生音を出力するもので、残響音や拍手、歓声に包まれた印象を与える効果がある。これにより、より立体的な音像を作り出すことができる。

【0024】図1に戻って、本実施例の多層符号化装置について説明する。この図1において、各入力端子201₁~201_nには、それぞれ対応する各チャンネルのデジタルオーディオ信号が供給される。

【0025】各入力端子201₁~201_nに供給された各チャンネルのデジタルオーディオ信号は、対応する各圧縮符号化回路202₁~202_nに送られて、ここで各チャンネル毎にそれぞれ圧縮符号化される。なお、各圧縮符号化回路202₁~202_nの詳細な構成については後述する。

【0026】各圧縮符号化回路202₁~202_nで圧縮符号化されたデジタルオーディオ信号は、メイン情報としてエントロピ符号化回路203₁~203_nに送られ、ここで、それぞれが可変長符号化される。また、上記圧縮符号化回路202₁~202_nでのデジタルオーディオ信号の圧縮符号化に関連する後述する語長情報やスケールファクタの情報は、サブ情報としてサブ情報圧縮回路204₁~204_nに送られ、ここで各チャンネル毎のビット配分情報に基づいて圧縮される。

【0027】上記エントロピ符号化回路203₁~203_nからのエントロピ符号化された各メイン情報は、適応量子化回路205₁~205_nに送られると共に、上記Logスペクトラルエンベロープ検出回路208に

も送られる。なお、図1では、チャンネルCH2～CHnの各エントロピ符号化回路203₂～203_nからのエントロピ符号化されたメイン情報は、それぞれ対応する端子210₂～210_nから、図示を省略している各チャンネルCH2～CHnに対応する適応量子化回路205₂～205_nに送られると共に、端子215₂～215_nを介して上記Logスペクトラルエンベロープ検出回路208に送られるようになっている。

【0028】上記エントロピ符号化された各メイン情報が供給されるLogスペクトラルエンベロープ検出回路208では、これら各チャンネルのメイン情報からそれぞれエネルギーを検出して、各チャンネルのLogスペクトラルエンベロープを検出する。当該各チャンネルのLogスペクトラルエンベロープ情報は、上記配分決定回路209に送られる。

【0029】当該配分決定回路209では、上記Logスペクトラルエンベロープ検出回路208からの各チャンネルのLogスペクトラルエンベロープ情報に基づいて、各チャンネルのメイン情報に対するチャンネル間ビット配分情報と、各サブ情報に対するチャンネル間ビット配分情報を決定する。

【0030】上記配分決定回路209からの上記メイン情報に対するチャンネル間ビット配分情報は、適応量子化回路205₁～205_nに送られ、上記サブ情報に対するチャンネル間ビット配分情報は、上記サブ情報圧縮回路204₁～204_nに送られる。なお、図1の例では、チャンネルCH2～CHnのメイン情報に対するチャンネル間ビット配分情報は端子214₂～214_nを介して図示を省略している適応量子化回路205₂～205_nに送られ、チャンネルCH2～CHnのサブ情報に対するチャンネル間ビット配分情報は端子217₂～217_n及び端子216₂～216_nを介してサブ情報圧縮回路204₂～204_nに送られる。

【0031】上記適応量子化回路205₁～205_nでは、上記メイン情報に対するチャンネル間ビット配分情報を用いて、上記各エントロピ符号化回路203₁～203_nからのエントロピ符号化されたデジタルオーディオ信号を適応的に量子化する。

【0032】また、上記サブ情報圧縮回路205₁～205_nでは、上記サブ情報に対するチャンネル間ビット配分情報を用いて、上記圧縮符号化回路202₁～202_nからのサブ情報（語長情報、スケールファクタ情報）を圧縮符号化する。

【0033】上記適応量子化回路205₁～205_nとサブ情報圧縮回路205₁～205_nからの各出力は、マルチプレクサ206に送られる。なお、チャンネルCH2～CHnの図示を省略している適応量子化回路205₂～205_nからの各出力は、端子212₂～212_nを介して上記マルチプレクサ206に送られ、チャンネルCH2～CHnのサブ情報圧縮回路205₂～205_nからの各出力は、端子213₂～213_nを介して上記マルチプレクサ206に送られる。

5_nからの各出力は、端子213₂～213_nを介して上記マルチプレクサ206に送られる。

【0034】当該マルチプレクサ206では、供給された各データをマルチプレクスして端子207から出力する。この出力データが符号化データとして例えば記録媒体等に記録される。

【0035】また、図1の本実施例装置においては、例えば図3に示すように、上記圧縮符号化回路202₁～202_nに対して供給する各チャンネルのデジタルオーディオ信号から、少なくとも2チャンネル毎（図3の例では全チャンネル毎）に対応してそれぞれ設けたハイパスフィルタ222によって高域成分のみを取り出した後、これら各高域成分を加算回路223によって加算し、端子225を介した当該高域成分の加算データに対して、所定の処理、例えば圧縮符号化、エントロピ符号化及び量子化の処理を施すようにすることもできる。このとき、上記高域成分を取り出す各チャンネルの入力デジタルオーディオ信号からは、それぞれローパスフィルタ221によって低域成分も取り出し、端子224を介して対応する圧縮符号化回路202₁～202_nに送るようにする。その後、上記高域及び低域成分を取り出す各チャンネルのうち、ある1つのチャンネル（図3の例ではチャンネルCH1）の上記マルチプレクサ206へ入力するメイン情報に対して、上記所定の処理を施した加算データを加算する処理（チャンネル間クロストーク処理）を行うようにする。

【0036】ここで、上記各チャンネルの高域成分を加算して、ある1つのチャンネルの低域成分に加算することを行うのは、以下の理由による。

【0037】すなわち、人間の耳は高域の成分に対する定位感が少なく、このため高域成分については例えば複数個のスピーカのうちのいずれか1つのスピーカからでても人間にはどのスピーカから出てきているのか聞き取り難いという性質がある。このため、複数チャンネルのオーディオ信号の各高域成分を、そのうちの例えば1チャンネルに対応するスピーカのみに送るようにしても、人間には各チャンネルのオーディオ信号の高域成分が当該1つのスピーカのみから出力されているとは感じられない。したがって、上述のように、上記各チャンネルの高域成分の音声を加算して、この加算データのある1つのチャンネルの低域成分に加算することで、各チャンネルの高域成分を1つのチャンネル分に圧縮できることになる。

【0038】次に、上記図1に示した圧縮符号化回路202₁～202_nの具体的な構成について説明する。

【0039】これら圧縮符号化回路202₁～202_nでは、オーディオPCM信号等の入力デジタル信号を、帯域分割符号化（SBC）、適応変換符号化（ATC）及び適応ビット配分（APC-AB）の各技術を用いて圧縮符号化している。

【0040】以下、図4を用いて説明する。図4に示す

本実施例の圧縮符号化回路 $202_1 \sim 202_n$ では、入力デジタル信号をフィルタなどにより複数の周波数帯域に分割すると共に、各周波数帯域毎に直交変換を行って、得られた周波数軸のスペクトルデータを、後述する人間の聴覚特性を考慮したいわゆる臨界帯域幅（クリティカルバンド）毎に適応的にビット配分して符号化している。この時、高域では臨界帯域幅を更に分割した帯域を用いる。もちろんフィルタなどによる非ブロッキングの周波数分割幅は等分割幅としてもよい。さらに、本実施例においては、直交変換の前に入力信号に応じて適応的にブロックサイズ（ブロック長）を変化させると共に、クリティカルバンド単位もしくは高域では臨界帯域幅（クリティカルバンド）を更に細分化したブロックでフローティング処理を行っている。このクリティカルバンドとは、人間の聴覚特性を考慮して分割された周波数帯域であり、ある純音の周波数近傍の同じ強さの狭帯域バンドノイズによって当該純音がマスクされるときそのノイズの持つ帯域のことである。このクリティカルバンドは、高域ほど帯域幅が広くなっており、上記 $0 \sim 22\text{ kHz}$ の全周波数帯域は例えば25のクリティカルバンドに分割されている。

【0041】すなわち、図4において、入力端子10には例えば $0 \sim 22\text{ kHz}$ のオーディオPCM信号が供給されている。この入力信号は、例えばいわゆるQMF等の帯域分割フィルタ11により $0 \sim 11\text{ kHz}$ 帯域と $11\text{ kHz} \sim 22\text{ kHz}$ 帯域とに分割され、 $0 \sim 11\text{ kHz}$ 帯域の信号は同じくいわゆるQMF等の帯域分割フィルタ12により $0 \sim 5.5\text{ kHz}$ 帯域と $5.5\text{ kHz} \sim 11\text{ kHz}$ 帯域とに分割される。帯域分割フィルタ11からの $11\text{ kHz} \sim 22\text{ kHz}$ 帯域の信号は、直交変換回路の一例であるMDCT（Modified Discrete Cosine Transform）回路13に送られ、帯域分割フィルタ12からの $5.5\text{ kHz} \sim 11\text{ kHz}$ 帯域の信号はMDCT回路14に送られ、帯域分割フィルタ12からの $0 \sim 5.5\text{ kHz}$ 帯域の信号はMDCT回路15に送られることにより、それぞれMDCT処理される。なお、各MDCT回路13、14、15では、各帯域毎に設けたブロック決定回路19、20、21により決定されたブロックサイズに基づいてMDCT処理がなされる。

【0042】ここで、上記ブロック決定回路19、20、21により決定される各MDCT回路13、14、15でのブロックサイズの具体例を図5のA及びBに示す。なお、図5のAには直交変換ブロックサイズが長い場合（ロングモードにおける直交変換ブロックサイズ）を、図5のBには直交変換ブロックサイズが短い場合（ショートモードにおける直交変換ブロックサイズ）を示している。この図5の具体例においては、3つのフィルタ出力は、それぞれ2つの直交変換ブロックサイズを持つ。すなわち、低域側の $0 \sim 5.5\text{ kHz}$ 帯域の信

は、長いブロック長の場合（図5のA）は1ブロック内のサンプル数を128サンプルとし、短いブロックが選ばれた場合（図5のB）には1ブロック内のサンプル数を32サンプル毎のブロックとしている。これに対して高域側の $11\text{ kHz} \sim 22\text{ kHz}$ 帯域の信号に対しては、長いブロック長の場合（図5のA）は1ブロック内のサンプル数を256サンプルとし、短いブロックが選ばれた場合（図5のB）には1ブロック内のサンプル数を32サンプル毎のブロックとしている。このようにして短いブロックが選ばれた場合には各帯域の直交変換ブロックのサンプル数を同じとして高域程時間分解能を上げ、なおかつブロック化に使用するウィンドウの種類を減らしている。なお、上記ブロック決定回路19、20、21で決定されたブロックサイズを示す情報は、後述の適応ビット配分符号化回路16、17、18に送られると共に、出力端子23、25、27から出力される。

【0043】再び図5において、各MDCT回路13、14、15にてMDCT処理されて得られた周波数領域のスペクトルデータあるいはMDCT係数データは、いわゆる臨界帯域（クリティカルバンド）または高域では更にクリティカルバンドを分割した帯域毎にまとめられて適応ビット配分符号化回路16、17、18に送られている。

【0044】適応ビット配分符号化回路16、17、18では、上記ブロックサイズの情報、及び臨界帯域（クリティカルバンド）または高域では更にクリティカルバンドを分割した帯域毎に割り当てられたビット数に応じて各スペクトルデータ（あるいはMDCT係数データ）を再量子化（正規化して量子化）するようにしている。適応ビット配分符号化回路16、17、18で符号化されたデータは、出力端子22、24、26を介して取り出される。また、当該適応ビット配分符号化回路16、17、18では、どのような信号の大きさに関する正規化がなされたかを示すスケールファクタ情報と、どのようなビット長で量子化がされたかを示すビット長情報（語長情報）も求めており、これらも同時に出力端子22、24、26から出力される。

【0045】また、図5における各MDCT回路13、14、15の出力からは、上記臨界帯域（クリティカルバンド）または高域では更にクリティカルバンドを分割した帯域毎のエネルギーが、例えば当該バンド内での各振幅値の2乗平均の平方根を計算すること等により求められる。もちろん、上記スケールファクタそのものを以後のビット配分の為に用いるようにしてもよい。この場合には新たなエネルギー計算の演算が不要となるため、ハード規模の節約となる。また、各バンド毎のエネルギーの代わりに、振幅値のピーク値、平均値等を用いることも可能である。

【0046】ここで、上記適応ビット配分回路16、17、18のより具体的な構成を図6で説明する。この図

6に示す適応ビット配分回路では、MDC T係数の大きさが各ブロック毎に求められ、そのMDC T係数が入力端子801に供給される。当該入力端子801に供給されたMDC T係数は、帯域毎のエネルギー算出回路803に与えられる。帯域毎のエネルギー算出回路803では、クリティカルバンド又は高域においてはクリティカルバンドを更に再分割したそれぞれの帯域に関する信号エネルギーを算出する。帯域毎のエネルギー算出回路803で算出されたそれぞれの帯域に関するエネルギーは、エネルギー依存ビット配分回路804に供給される。

【0047】エネルギー依存ビット配分回路804では、使用可能総ビット発生回路802からの使用可能総ビット、本実施例では128Kbpsの内のある割合（本実施例では100Kbps）を用いて白色の量子化雑音を作り出すようなビット配分を行う。このとき、入力信号のトナリティが高いほど、すなわち入力信号のスペクトルの凸凹が大きいほど、このビット量が上記128Kbpsに占める割合が増加する。なお、入力信号のスペクトルの凸凹を検出するには、隣接するブロックのブロックフローティング係数の差の絶対値の和を指標として使う。そして、求められた使用可能なビット量につき、各帯域のエネルギーの対数値に比例したビット配分を行う。

【0048】聴覚許容雑音レベルに依存したビット配分算出回路805は、先ず上記クリティカルバンド毎に分割されたスペクトルデータに基づき、いわゆるマスキング効果等を考慮した各クリティカルバンド毎の許容ノイズ量を求め、次に聴覚許容雑音スペクトルを与えるように使用可能総ビットからエネルギー依存ビットを引いたビット分が配分される。このようにして求められたエネルギー依存ビットと聴覚許容雑音レベルに依存したビットは加算されて、図4の適応ビット配分符号化回路16、17、18により各クリティカルバンド毎もしくは高域においてはクリティカルバンドを更に複数帯域に分割した帯域に割り当てられたビット数に応じて各スペクトルデータ（あるいはMDC T係数データ）を再量子化するようにしている。このようにして符号化されたデータは、図4の出力端子22、24、26を介して取り出される。

【0049】さらに詳しく上記聴覚許容雑音スペクトル依存のビット配分回路805中の聴覚許容雑音スペクトル算出回路について説明すると、MDC T回路13、14、15で得られたMDC T係数が上記許容雑音算出回路に与えられる。

【0050】図7は上記許容雑音算出回路をまとめて説明した一具体例の概略構成を示すブロック回路図である。この図7において、入力端子521には、MDC T回路13、14、15からの周波数領域のスペクトルデータが供給されている。

エネルギー算出回路522に送られて、上記クリティカルバンド（臨界帯域）毎のエネルギーが、例えば当該バンド内での各振幅値2乗の総和を計算すること等により求められる。この各バンド毎のエネルギーの代わりに、振幅値のピーク値、平均値等が用いられることもある。このエネルギー算出回路522からの出力として、例えば各バンドの総和値のスペクトルは、一般にバークスペクトルと称されている。図8はこのような各クリティカルバンド毎のバークスペクトルSBを示している。ただし、この図8では、図示を簡略化するため、上記クリティカルバンドのバンド数を12バンド（B1～B12）で表現している。

【0052】ここで、上記バークスペクトルSBのいわゆるマスキングに於ける影響を考慮するために、該バークスペクトルSBに所定の重み付け関数を掛けて加算するような畳込み（コンボリューション）処理を施す。このため、上記帯域毎のエネルギー算出回路522の出力すなわち該バークスペクトルSBの各値は、畳込みフィルタ回路523に送られる。該畳込みフィルタ回路523は、例えば、入力データを順次遅延させる複数の遅延素子と、これら遅延素子からの出力にフィルタ係数（重み付け関数）を乗算する複数の乗算器（例えば各バンドに対応する25個の乗算器）と、各乗算器出力の総和をとる総和加算器とから構成されるものである。なお、上記マスキングとは、人間の聴覚上の特性により、ある信号によって他の信号がマスクされて聞こえなくなる現象をいうものであり、このマスキング効果には、時間領域のオーディオ信号による時間軸マスキング効果と、周波数領域の信号による同時刻マスキング効果とがある。これらのマスキング効果により、マスキングされる部分にノイズがあったとしても、このノイズは聞こえないことになる。このため、実際のオーディオ信号では、このマスキングされる範囲内のノイズは許容可能なノイズとされる。

【0053】ここで、上記畳込みフィルタ回路523の各乗算器の乗算係数（フィルタ係数）の一具体例を示すと、任意のバンドに対応する乗算器Mの係数を1とするとき、乗算器M-1で係数0.15を、乗算器M-2で係数0.0019を、乗算器M-3で係数0.0000086を、乗算器M+1で係数0.4を、乗算器M+2で係数0.06を、乗算器M+3で係数0.007を各遅延素子の出力に乗算することにより、上記バークスペクトルSBの畳込み処理が行われる。ただし、Mは1～25の任意の整数である。

【0054】次に、上記畳込みフィルタ回路523の出力は引算器524に送られる。該引算器524は、上記畳込んだ領域での後述する許容可能なノイズレベルに対応するレベル α を求めるものである。なお、当該許容可能なノイズレベル（許容ノイズレベル）に対応するレベル α は、後述するように、適応マスキング処理を

行うことによって、クリティカルバンドの各バンド毎の許容ノイズレベルとなるようなレベルである。ここで、上記引算器 524 には、上記レベル α を求めるための許容関数（マスキングレベルを表現する関数）が供給される。この許容関数を増減させることで上記レベル α の制御を行っている。当該許容関数は、次に説明するような $(n - a_i)$ 関数発生回路 525 から供給されているものである。

【0055】すなわち、許容ノイズレベルに対応するレベル α は、クリティカルバンドのバンドの低域から順に与えられる番号を i とすると、次の式で求めることができる。

$$\alpha = S - (n - a_i)$$

この式において、 n 、 a は定数で $a > 0$ 、 S は畳込み処理されたバークスペクトルの強度であり、式中 $(n - a_i)$ が許容関数となる。例として $n = 38$ 、 $a = -0.5$ を用いることができる。

【0056】このようにして、上記レベル α が求められ、このデータは、割算器 526 に伝送される。当該割算器 526 では、上記畳込みされた領域での上記レベル α を逆コンボリューションするためのものである。したがって、この逆コンボリューション処理を行うことにより、上記レベル α からマスキングスレッシュホールドが得られるようになる。すなわち、このマスキングスレッシュホールドが許容ノイズスペクトルとなる。なお、上記逆コンボリューション処理は、複雑な演算を必要とするが、本実施例では簡略化した割算器 526 を用いて逆コンボリューションを行っている。

【0057】次に、上記マスキングスレッシュホールドは、合成回路 527 を介して減算器 528 に伝送される。ここで、当該減算器 528 には、上記帯域毎のエネルギー検出回路 522 からの出力、すなわち前述したバークスペクトル S_B が、遅延回路 529 を介して供給されている。したがって、この減算器 528 で上記マスキングスレッシュホールドとバークスペクトル S_B との減算演算が行われることで、図 3 に示すように、上記バークスペクトル S_B は、該マスキングスレッシュホールド MS のレベルで示すレベル以下がマスキングされることになる。なお、遅延回路 529 は上記合成回路 527 以前の各回路での遅延量を考慮してエネルギー検出回路 522 からのバークスペクトル S_B を遅延させるために設けられている。

【0058】当該減算器 528 からの出力は、許容雑音補正回路 530 を介し、出力端子 531 を介して取り出され、例えば配分ビット数情報が予め記憶された ROM 等（図示せず）に送られる。この ROM 等は、上記減算回路 528 から許容雑音補正回路 530 を介して得られた出力（上記各バンドのエネルギーと上記ノイズレベル設定手段の出力との差分のレベル）に応じ、各バンド毎の

【0059】このようにしてエネルギー依存ビットと聴覚許容雑音レベルに依存したビットは加算されてその配分ビット数情報を用いて当該適応ビット配分符号化回路 16、17、18 では符号化が行われる。

【0060】すなわち要約すれば、適応ビット配分符号化回路 16、17、18 では、上記クリティカルバンドの各バンド帯域（クリティカルバンド）毎もしくは高域においてはクリティカルバンドを更に複数帯域に分割した帯域のエネルギーもしくはピーク値と上記ノイズレベル設定手段の出力との差分のレベルに応じて配分されたビット数で上記各バンド毎のスペクトルデータを量子化することになる。

【0061】ところで、上述した合成回路 527 での合成の際には、最小可聴カーブ発生回路 532 から供給される図 9 に示すような人間の聴覚特性であるいわゆる最小可聴カーブ RC を示すデータと、上記マスキングスレッシュホールド MS とを合成することができる。この最小可聴カーブにおいて、雑音絶対レベルがこの最小可聴カーブ以下ならば該雑音は聞こえないことになる。この最小可聴カーブは、コーディングが同じであっても例えば再生時の再生ボリュームの違いで異なるものとなが、現実的なデジタルシステムでは、例えば 16 ビットダイナミックレンジへの音楽のはいり方にはさほど違いがないので、例えば 4 kHz 付近の最も耳に聞こえやすい周波数帯域の量子化雑音が聞こえないとすれば、他の周波数帯域ではこの最小可聴カーブのレベル以下の量子化雑音は聞こえないと考えられる。したがって、このように例えばシステムの持つダイナミックレンジの 4 kHz 付近の雑音が聞こえない使い方をすると仮定し、この最小可聴カーブ RC とマスキングスレッシュホールド MS とを共に合成することで許容ノイズレベルを得るようにすると、この場合の許容ノイズレベルは、図 9 中の斜線で示す部分までとすることができるようになる。なお、本実施例では、上記最小可聴カーブの 4 kHz のレベルを、例えば 20 ビット相当の最低レベルに合わせている。また、この図 9 は、信号スペクトル S_S も同時に示している。

【0062】また、上記許容雑音補正回路 530 では、補正情報出力回路 533 から送られてくる例えば等ラウドネスカーブの情報に基づいて、上記減算器 528 からの出力における許容雑音レベルを補正している。ここで、等ラウドネスカーブとは、人間の聴覚特性に関する特性曲線であり、例えば 1 kHz の純音と同じ大きさに聞こえる各周波数での音の音圧を求めて曲線で結んだもので、ラウドネスの等感度曲線とも呼ばれる。またこの等ラウドネス曲線は、図 9 に示した最小可聴カーブ RC と略同じ曲線を描くものである。この等ラウドネス曲線においては、例えば 4 kHz 付近では 1 kHz のところより音圧が 8～10 dB 下がっても 1 kHz と同じ大き

よりも約 15 dB 高くないと同じ大きさに聞こえない。このため、上記最小可聴カーブのレベルを越えた雑音（許容ノイズレベル）は、該等ラウドネス曲線に応じたカーブで与えられる周波数特性を持つようにするのが良いことがわかる。このようなことから、上記等ラウドネス曲線を考慮して上記許容ノイズレベルを補正することは、人間の聴覚特性に適合していることがわかる。

【0063】以上述べた聴覚許容雑音レベルに依存したスペクトル形状を使用可能総ビット 128 Kbps の内のある割合を用いるビット配分でつくる。この割合は入力信号のトナリリティが高くなるほど減少する。

【0064】次にビット量分割手法について説明する。図 6 に戻って、MDC T 回路出力が供給される入力端子 801 からの信号は、スペクトルの滑らかさ算出回路 808 にも与えられ、ここでスペクトルの滑らかさが算出される。本実施例では、信号スペクトルの絶対値の隣接値間の差の絶対値の和を、信号スペクトルの絶対値の和で割った値を、上記スペクトルの滑らかさとして算出している。

【0065】上記スペクトルの滑らかさ算出回路 808 の出力は、ビット分割率決定回路 809 に与えられ、ここでエネルギー依存のビット配分と、聴覚許容雑音スペクトルによるビット配分間のビット分割率が決定される。ビット分割率はスペクトルの滑らかさ算出回路 808 の出力値が大きいほど、スペクトルの滑らかさが無いと考えると、エネルギー依存のビット配分よりも、聴覚許容雑音スペクトルによるビット配分に重点をおいたビット配分を行う。ビット分割率決定回路 809 は、それぞれエネルギー依存のビット配分及び聴覚許容雑音スペクトルによるビット配分の大きさをコントロールするマルチプライヤ 811 及び 812 に対してコントロール出力を送る。ここで、仮にスペクトルが滑らかであり、エネルギー依存のビット配分に重きをおくように、マルチプライヤ 811 へのビット分割率決定回路 809 の出力が 0.8 の値を取ったとき、マルチプライヤ 812 へのビット分割率決定回路 809 の出力は $1 - 0.8 = 0.2$ とする。これら 2 つのマルチプライヤの出力はアダー 806 で足し合わされて最終的なビット配分情報となって、出力端子 807 から出力される。

【0066】次に、図 10 には、上記図 4 の圧縮符号化回路に対応する伸張復号化回路の構成を示す。

【0067】この図 10 において、各帯域の量子化された MDC T 係数は、入力端子 122、124、126 に与えられ、使用されたブロックサイズ情報は、入力端子 123、125、127 に与えられる。復号化回路 116、117、118 では適応ビット配分情報を用いてビット割当を解除する。

【0068】次に、IMDC T 回路 113、114、115 では周波数領域の信号が時間領域の信号に変換され、その後、I Q M F 回路 111、112 で帯域合成がなされ、出力端子 110 から出力される。

112、111 により、全体域信号に復号化される。その後、I Q M F 回路 111、112 で帯域合成がなされ、出力端子 110 から出力される。

【0069】

【発明の効果】上述したように、本発明の多層符号化装置においては、第 1 の符号化手段で各チャネルのデジタルオーディオ信号を圧縮符号化した後、この第 1 の符号化手段の出力符号化信号に対して、さらに第 2 の符号化手段によってチャネル毎のエネルギーに略応じた比率で、各チャネルの情報を割り当てる符号化を行うことにより、チャネル間のビット配分と高圧縮率を可能としている。

【図面の簡単な説明】

【図 1】本発明実施例の多層符号化装置の概略構成を示すブロック回路図である。

【図 2】8 チャネルデジタルサラウンドシステムにおけるスピーカの配置を説明するための図である。

【図 3】チャネル間クロストークを行う主要構成要素の概略構成を示すブロック回路図である。

【図 4】本実施例の圧縮符号化回路の構成例を示すブロック回路図である。

【図 5】圧縮符号化回路での信号の周波数及び時間分割を示す図である。

【図 6】情報信号の大きさ及び聴覚許容雑音スペクトルの 2 者を用いたビット配分手法を実現する構成を示すブロック回路図である。

【図 7】許容雑音レベルを求める構成を示すブロック回路図である。

【図 8】各帯域の信号レベルによるマスキングスレシヨールドの例を示す図である。

【図 9】情報スペクトル、マスキングスレシヨールド、最小可聴限を示す図である。

【図 10】本実施例の圧縮符号化回路に対応する伸張復号化回路の構成例を示すブロック回路図である。

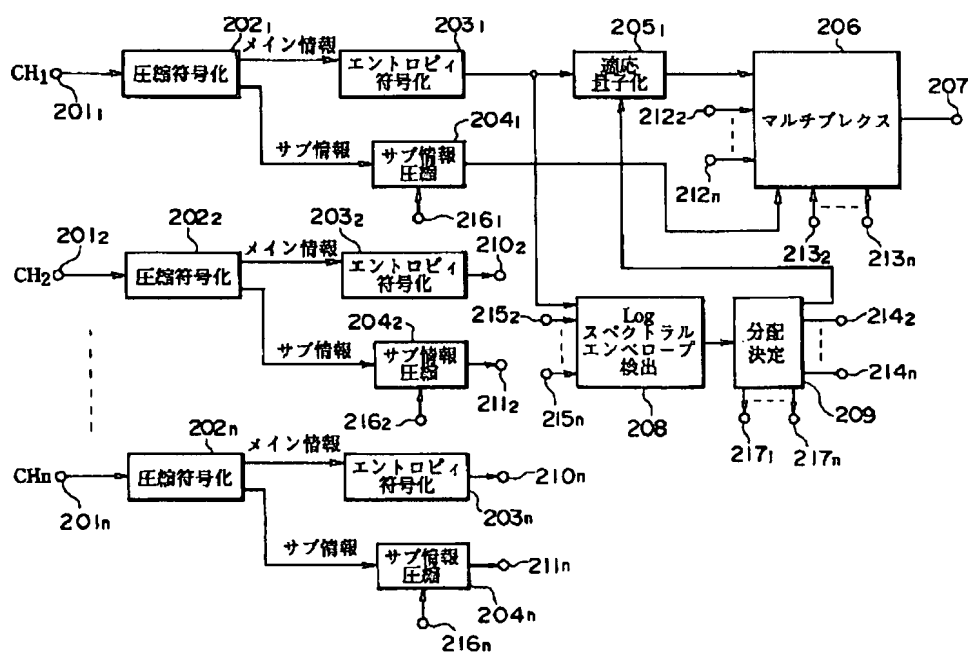
【符号の説明】

202₁ ~ 202_n . . . 圧縮符号化回路
 203₁ ~ 203_n . . . エントロピ符号化回路
 204₁ ~ 204_n . . . サブ情報圧縮回路
 205₁ ~ 205_n . . . 適応量子化回路
 206 . . . マルチプレクサ
 208 . . . L o g スペクトラルエンベロープ検出回路
 209 . . . 分配決定回路
 221 . . . ローパスフィルタ
 222 . . . ハイパスフィルタ
 223 . . . 加算回路
 11, 12 帯域分割フィルタ
 13, 14, 15 MDC T 回路
 16, 17, 18 適応ビット配分符号化回路
 19, 20, 21 ブロックサイズ決定回路
 22, 23, 24 復号化出力端子

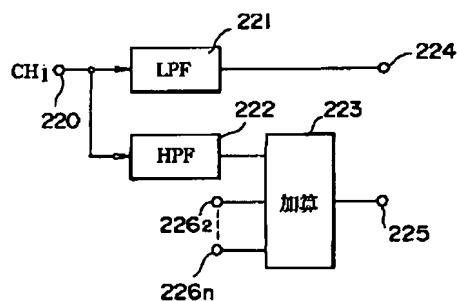
23, 25, 27・・・ブロックサイズ情報出力端子
 116, 117, 118・・・適応ビット配分復号化回路
 113, 114, 115・・・IMDCT回路
 112, 111・・・IQMF回路
 520・・・許容雑音算出回路
 521・・・許容雑音算出回路入力端子
 522・・・帯域毎のエネルギー検出回路
 523・・・畳込みフィルタ回路
 524・・・引算器
 525・・・ $n-a$ 関数発生回路
 526・・・割算器
 527・・・合成回路

528・・・減算器
 530・・・許容雑音補正回路
 532・・・最小可聴カーブ発生回路
 533・・・補正情報出力回路
 801・・・MDCT回路出力端子
 802・・・使用可能総ビット発生回路
 803・・・帯域毎のエネルギー算出回路
 804・・・エネルギー依存のビット配分回路
 805・・・聴覚許容雑音レベル依存のビット配分回路
 806・・・アダー
 807・・・各帯域のビット割当量出力端子
 808・・・スペクトルの滑らかさ算出回路
 809・・・ビット分割率決定回路
 811, 812・・・マルチプライヤ

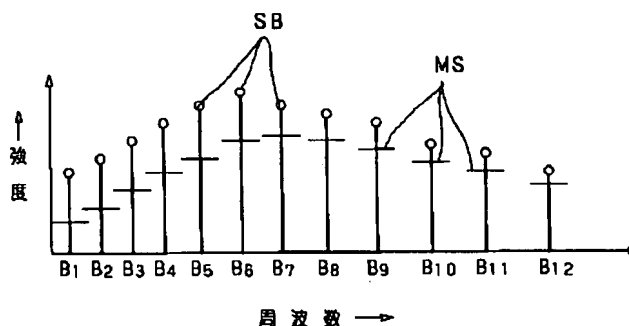
【図1】



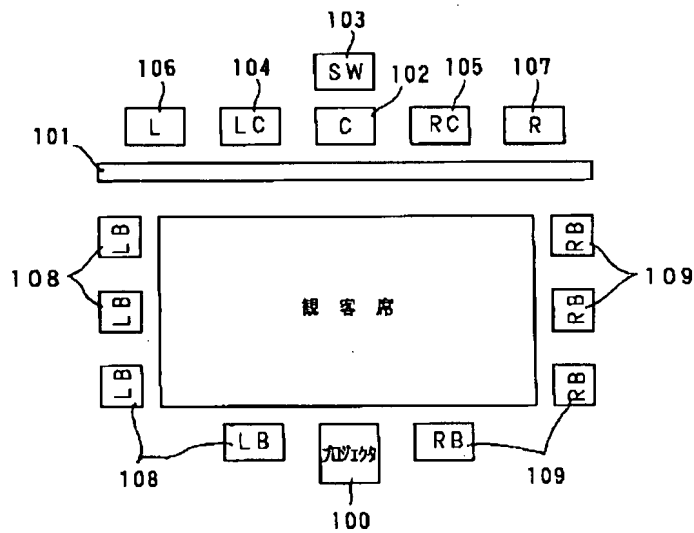
【図3】



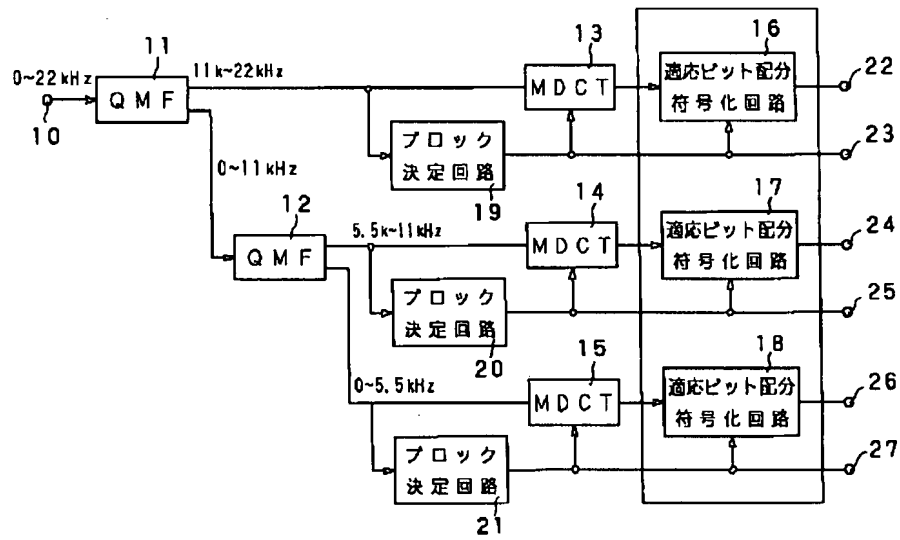
【図8】



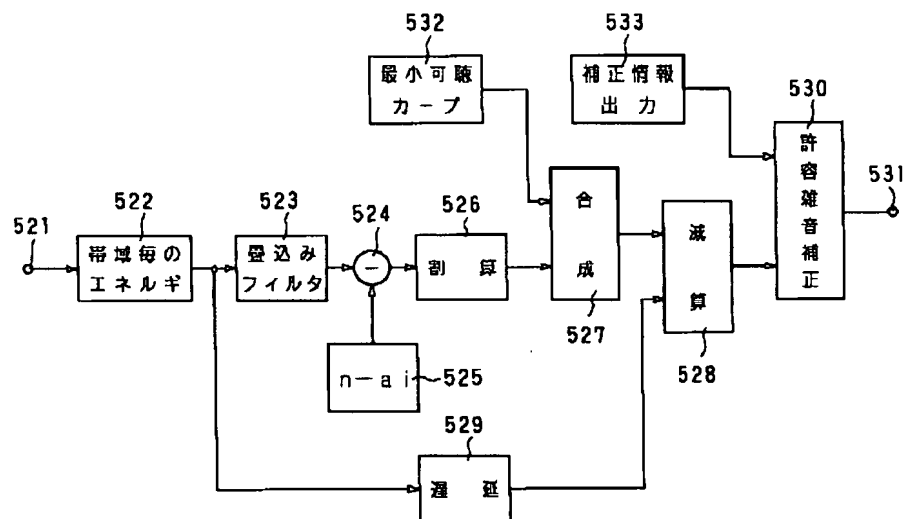
【図2】



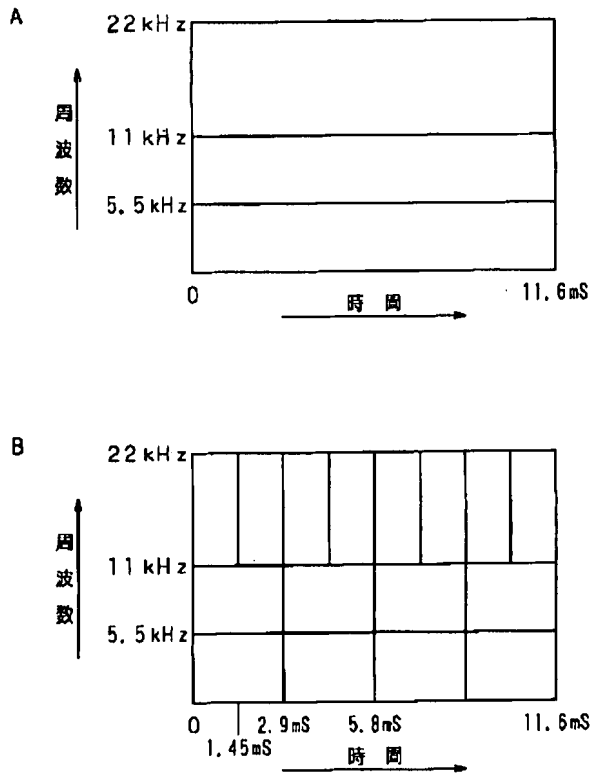
【図4】



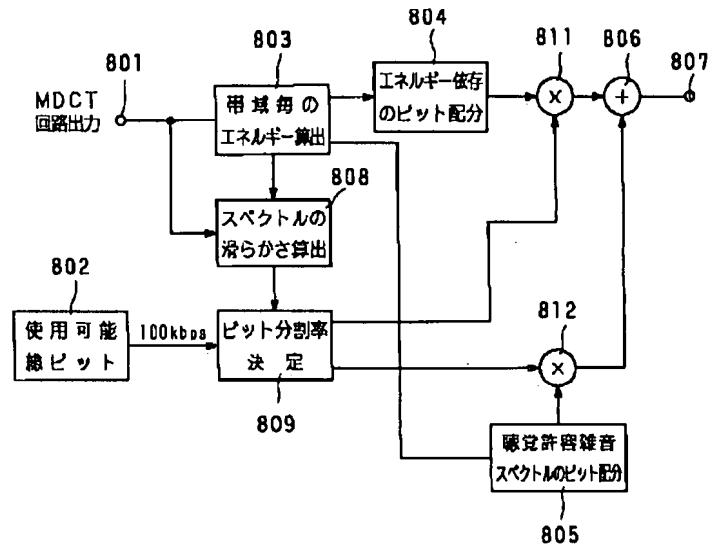
【図7】



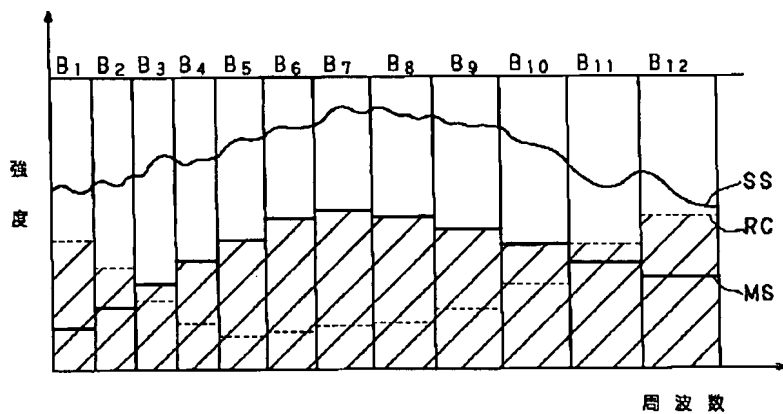
【図 5】



【図 6】



【図 9】



【図 10】

